(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 September 2003 (12.09.2003)

PCT

(10) International Publication Number
WO 03/074784 A1

(51) International Patent Classification⁷: D21G 9/00, G05B 13/00

(21) International Application Number: PCT/GB03/00943

(22) International Filing Date: 5 March 2003 (05.03.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0205336.1 7 March 2002 (07.03.2002) GB

(71) Applicant *(for all designated States except US)*: **UNIVERSITY OF MANCHESTER INSTITUTE OF SCIENCE & TECHNOLOGY** [GB/GB]; P.O. Box 88, Manchester M60 1QD (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **VIRTA, Marko, Tapani** [FI/FI]; Palojoentie 1 C 54, 16300 Orimattila (FI). **WANG, Hong** [GB/GB]; 17 Holbeach Close, Bury BL8 1XA (GB).

(74) Agent: ALLMAN, Peter, John; Marks & Clerk, Sussex House, 83-85 Mosley Street, Manchester M2 3LG (GB).

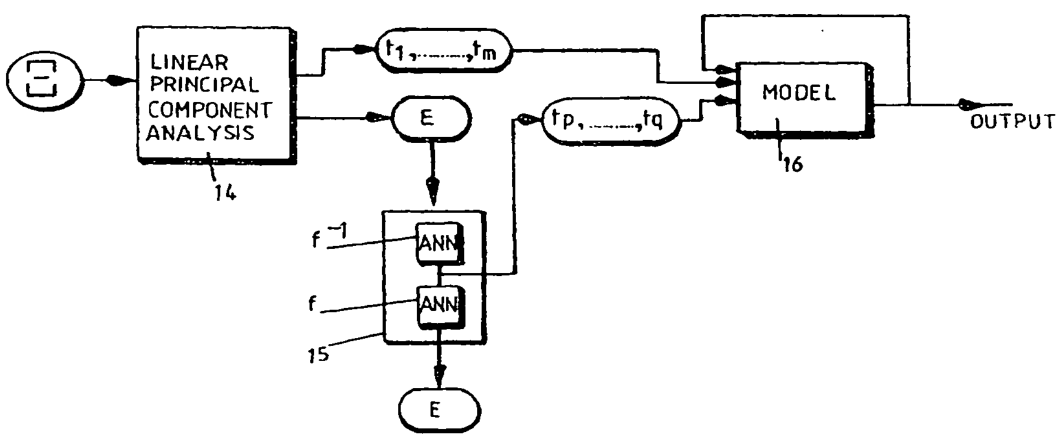
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

[Continued on next page]

(54) Title: A METHOD OF DYNAMICALLY MODELLING A PAPER MANUFACTURING PLANT USING PCA (PRINCIPAL COMPONENT ANALYSIS)



(57) Abstract: A method for converting a quantity of input data to output data of lesser dimension. A linear principal component analysis technique is applied to the input data to generate a plurality of linear principal components and an error signal. The error signal is input to a first neural network which outputs at least one variable, the said at least one variable is input to a second neural network and the first and second neural networks are configured such that the output of the second neural network is substantially equal to the input to the first neural network. The output data is represented by the said plurality of linear principal components and the said at least one variable. There is also provided a method of dynamically modelling a paper manufacturing plant comprising derivation of a function which takes as input a plurality of parameters of said paper manufacturing plant which have been reduced in dimension and outputs a value indicative of a quality of paper output from the paper manufacturing plant.

WO 03/074784 A1

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

A METHOD OF DYNAMICALLY MODELLING A PAPER MANUFACTURING PLANT USING PCA (PRINCIPAL COMPONENT ANALYSIS)

The present invention relates to a method for converting a quantity of input data to output data of lesser dimension. The present invention also relates to a method of modelling a paper manufacturing plant so as to allow optimisation of various parameters within the manufacturing process.

A paper manufacturing plant comprises two distinct phases - a first wet end and a second dry end. The present invention is concerned with parameters of the wet end process, which affect both phases of the paper manufacturing plant.

The wet end process involves input of fibres, water and chemicals. The wet end process comprises a number of complex processes including considerations of fluid dynamics, chemical reactions and physical reactions. This multi-input process is further complicated by various machine operating considerations such as speed of operations, chemical inputs and the particular configuration of the paper manufacturing machine

Currently, systems exist which allow monitoring of wet end inputs and allow alteration of these inputs in response to obtained measurements. One such system is a closed loop pH control system. This system measures pH within the head box and adjusts one or more inputs so as to ensure that the pH remains within a predetermined range in order to produce paper with predetermined properties. This closed loop system is not currently in widespread use. Additionally, the system provides no indication of likely changes in output qualities in response to these input changes.

Currently wet end parameters are adjusted according to an operator's individual experience and expertise using a "best guess" approach. While experienced operators may achieve good results using this method, it is likely that those with less experience will be unable to obtain satisfactory settings of these parameters, thereby yielding inconsistent results. Additionally, relying on such human intuition is never likely to

result in optimum paper manufacture parameters, no matter how much experience the operator may possess.

Recently, a wet end information centre (WIC) system has been employed in some paper manufacture plants in order to obtain a wide variety of data relating to various operating parameters of the wet end paper manufacture process. Additionally, data relating to various user specified parameters such as speed is readily available.

Despite having this large quantity of data available, use of this data as a basis for a model of the paper manufacturing process has not heretofore been considered. Such a model would ideally allow prediction of predetermined output characteristics in response to changes in predetermined input characteristics. However, given the quantity of input data and the complex permutations of this data that must be considered to derive optimum parameter values, a model cannot be created and optimised using available computing power.

It is an object of the present invention to obviate or mitigate one or more of the problems outlined above.

According to a first aspect of the present invention there is provided a method for converting a quantity of input data to output data of lesser dimension, comprising:

applying a linear principal component analysis technique to the input data to generate a plurality of linear principal components and an error signal;.

inputting said error signal to a first neural network which outputs at least one variable;

inputting the said at least one variable to a second neural network; and

configuring the first and second neural networks such that the output of the second neural network is substantially equal to the input to the first neural network;

the output being represented by the said plurality of linear principal components and the said at least one variable.

Preferably, the first and second neural networks are trained until the difference between the output of the second neural network and the input to the first neural network is less than a predetermined threshold. Furthermore, neurons may be added to the first and second neural networks until the difference between the output of the second neural network and the input to the first neural network is less than a predetermined threshold. A difference between the input to the first neural network and output of the second neural network may be assessed using an auto correlation technique.

Preferably, the first and second neural networks each comprise an input layer having connections with a hidden layer which in turn has connections with an output layer, such that the output layer of the first neural network forms the input layer of the second neural network. Preferably, the hidden layer of each neural network has more neurons than the input layer of the first neural network, and more preferably, the hidden layer of each neural network has twice as many neurons as the input layer of the first neural network. The input layer of the first neural network may have an equal number of neurons to the output layer of the second neural network.

The method may further comprise using the output data as input to a model of a process, such that the model outputs at least one value representing a property of that process, comparing the value representing the said property output from the model with a measured value for that property, and training the model so as to adjust weights of nodes within the model until the difference between the value for the property output from the model and the measured value of the property is less than a predetermined threshold. The model may be implemented by means of a neural network.

According to a second aspect of the present invention, there is provided a method of dynamically modelling a paper manufacturing plant comprising derivation of a function which takes as input a plurality of parameters of said paper manufacturing plant which have been reduced in dimension and outputs a value indicative of a quality of paper output from the paper manufacturing plant.

Preferably, the plurality of parameters comprise a plurality of temporally spaced values for at least one parameter of the paper manufacturing plant. The function may also take as input values previously output from the function. The function may be provided by an artificial neural network.

At least some of the plurality of parameters may be output from a method according to the first aspect of the present invention.

The model may have a general form:

$$R_{k+d} = \breve{g}(S_k, S_{k-1}, S_{k-2}, \ldots, S_{k-(l-1)}, R_k, R_{k-1}, R_{k-2}, \ldots, R_{k-(l-1)}, t_{1(k)}, t_{2,(k)}, t_{3(k)}, t_{4(k)}, U_{(k)})$$

where: R represents a quality of the output paper;

k is a current sample;

k-1 is a previous sample;

$t_1, \ldots t_4$ are parameters of the paper manufacturing plant

$U_{(k-d)}$ is a variable related to chemical costs;

d is time delay; and

l is the time period over which previous measurements are to be taken into account.

A plurality of models are preferably generated and combined to generate a performance function. Optimisation of the performance function may seek to maximise at least one of the outputs of the plurality of models. Optimisation of the performance function may seek to focus at least one of the outputs of the plurality of models upon a predetermined target.

The performance function may have a form:

$$J = \frac{a}{A} + \frac{b}{B} + C$$

where: J is an output from the function;

A and B are outputs from models generated using a method according to any one of claims 1 to 10;

C is a variable representing all adjustable inputs that are to be minimised; and

a and b are constants.

Alternatively, the performance function may have a form:

$$J = a(A-T)^2 + \frac{b}{B} + C$$

where: J is an output from the function;

A and B are outputs from models generated using a method according to any one of claims 1 to 10;

T is a target upon which A is to be focussed; and

a and b are constants.

Alternatively, the performance function may have a form:

$$J = a(A-T)^2 + b(B-T_2)^2 + C$$

where: J is an output from the function;

A and B are outputs from models generated using a method according to any one of claims 1 to 10;

T and T_2 are target values for A and B respectively; and

a and b are constants.

Any of the performance functions detailed above may include at least one other term.

An embodiments of the present invention will now be described, by way of example, with reference to the accompanying drawings, in which

Figure 1 is a schematic illustration of a known paper manufacturing plant;

Figure 2 is a schematic illustration of inputs to and outputs from the wet end paper manufacturing process operated in the plant of figure 1;

Figure 3 is a flow chart of a method in accordance with the present invention;

Figure 4 is a schematic illustration of two Artificial Neural Networks used in a preferred embodiment of the present invention; and

Figure 5 is a graph showing results obtained using an autocorrelation method to compare input to and output from the networks of Figure 4.

Referring to figure 1, a complete paper manufacturing installation is illustrated in outline. The installation has a plurality of inputs, that is chemicals (represented by an arrow 1), fibres (represented by an arrow 2) and water (represented by an arrow 3). These inputs are all combined by a mixer 4 to form a raw pulp which is fed into a head box 5. The head box 5 feeds the pulp at a controlled rate onto a moving wire table 6 upon which the paper is formed. The wire table 6 comprises a number of vacuum boxes 7 which exert a force on the paper so as to extract from it as much water as possible. Water removed from paper on the wire table 6 by the vacuum boxes 7 is fed to a white water tank 8.

Water reaching the white water tank 8 can be recycled by input to the mixer unit 4 at a later stage. It will be appreciated that the white water must be reasonably pure in order to be effectively recycled. The quantity of chemicals reaching the white water tank should, where possible, be minimised.

Paper passing over the wire table 6 is kept flat by movement of a further rotatable wire table 9 which is positioned above the wire table 6. The space between these two wire tables 6, 9 is adjusted such that paper of desired thickness is manufactured.

The elements of the manufacturing process as described above make up the wet end process and it is this section of the manufacturing plant that the present invention seeks to optimise by use of modelling techniques. The remainder of the process, known as the dry end, will now be described for the sake of completeness.

Once formed on the wire table 6, the paper is passed through a press section, comprising four rollers 10. The rollers 10 exert considerable pressure on the paper so

as to remove as much water as possible by a squeezing action. Having passed through the press section the paper passes over a number of rollers 11 in a drying section where a squeezing action is continued and is enhanced by the effects of heat so as to remove water by evaporation.

The final stage of the process is known as a calendar section. Here, the paper is passed around a number of rollers 12 so as to obtain a paper having desired properties. For example, the calendar section may polish the paper so as to remove any roughness. Paper leaving the calendar section is rolled about a roller 13 which is the end process of the section. Those skilled in the art will realise that some paper manufacturing plants do not include a calendar section.

As described above, a number of inputs are mixed together to form the raw pulp, namely chemicals, water and fibres. The correct balance of these individual components is essential in order to obtain a finished product having desired properties. The paper plant may use a WIC system in order to obtain information about various operating parameters in the wet end of the manufacturing process. One known WIC system collects twenty four variables classified into four groups, each group having six measurements *viz* pH, temperature, turbidity, cationic demand, alkalinity and conductivity. The various groups are distributed across the wet end of the paper manufacture process. The measurement cycle provided is relatively slow, taking of the order of 15 minutes to obtain a complete set of measurements. This WIC data is combined with other user defined measurements such as wire table speed and pulp flow rate, and details of chemical properties of raw materials that are input to the process.

The combined data amounts to some eighty four variables and these variables offer a complete overview of the wet end manufacture process. Additionally, these variables are directly related to properties of the paper that is output from the manufacturing process such as strength and retention of chemical inputs. This is schematically illustrated in figure 2 where the three inputs to the process are shown on the left hand side of the diagram, the process is schematically represented in the centre and output

characteristics are represented on the right hand side of the diagram. The present invention provides a model relating these inputs to these outputs.

The model relies upon artificial neural networks (ANNs). Artificial neural networks are well known as is their use in modelling to provide an output prediction based on a plurality of input variables. The networks are trained using historical data such that the output accurately represents the real world response in such circumstances.

Typically, an ANN has the form:

$$y = \overline{\theta}^T \operatorname{tang}\left(\overline{\varpi}^T \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + \overline{B}_1 \right) + B_2 \qquad (1)$$

where:

y is a model output;

$\begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ is a vector containing n input variables;

$\overline{\theta}^T, \overline{\varpi}^T, \overline{B}_1, \overline{B}_2$ are vectors containing weights for each of the n input variables, ^{T} denoting that the vector has been transposed; and

tang is the standard sigmoid or hyperbolic tangent function.

Usually, a neural network is trained so as to determine optimum weights, before inputs are used to predict an output. In a preferred embodiment of the present invention, ANNs are trained using an approach similar to that described in R. Noreiga and H. Wang: "A direct adaptive neural network control for unknown non-linear systems, and its application", IEEE Transactions on Neural Networks, Volume 9, pp 27-34, 1998.

It was mentioned above that data collected from various parts of the wet end paper manufacturing process amounts to a large number of variables. It is not possible to use this quantity of data as input to an ANN and achieve a reasonable result within a

reasonable time. Therefore, techniques are required to reduce the dimension of the collected data.

Input data to a neural network can be considered to comprise n variables each having values for m time periods. That is a set of data $\{x_1, x_2, \ldots, x_n\}$ which can be represented by Ξ and is in fact:

$$\Xi = \begin{bmatrix} x_{11} x_{21} \cdots\cdots\cdots x_{n1} \\ \vdots \quad \vdots \quad \ddots \quad \vdots \\ \vdots \quad \vdots \quad \ddots \quad \vdots \\ \vdots \quad \vdots \quad \ddots \quad \vdots \\ x_{1m} x_{2m} \cdots\cdots\cdots x_{nm} \end{bmatrix} \qquad (2)$$

The present invention employs principal component analysis techniques to reduce the dimension of the collected data. That is, all data is input to a first process which reduces the dimension of the collected data. Linear and non-linear principal component analysis techniques are well known. In accordance with the present invention, the two techniques are combined to achieve the required reduction in data dimension.

The process is illustrated in figure 3. The input data Ξ is passed through a linear principal component analysis module 14 to generate a plurality of linear principal components $t_1,\ldots,t_m$ and an error signal E. The generated error signal E is fed into a non-linear principal component analysis module 15 which contains two ANNs denoted f^1 and f. An output from the non-linear principal component analysis module 15 comprises one or more non-linear principal components $t_p,\ldots,t_q$. The linear and non-linear principal component modules form input to a model 16. The function of the linear and non-linear principal component analysis modules is described in further detail below

The linear principal component analysis technique takes the data set Ξ as input, and outputs a plurality of linear principal components and an error signal. The operation of this linear technique is described below.

Given the input data set Ξ as in equation (2), a matrix Λ may be formed such that

$$\Lambda = \Xi^{T}\Xi \tag{3}$$

where

Ξ^{T} is the transpose of Ξ.

Eigenvalues can be computed for Λ to give a set of values:

$$\lambda_1 > \lambda_2 > \lambda_3 > \ldots > \lambda_n \tag{4}$$

where

$\lambda_1 \ldots \lambda_n$ are Eigenvalues; and

n is the dimension of Λ.

The original dataset Ξ can be expressed in terms of a number of components:

$$\Xi = t_1 p_1 + t_2 p_2 + \ldots + t_N p_N + E_0 \tag{5}$$

where

t_i is the principal component related to λ_i and is calculated as presented below;

p_i is the transpose of the Eigenvector of the corresponding Eigenvalue λ_i; and

E_0 is an error signal;

Determination of t_1 will be presented by way of example. t_1 is calculated by taking equation (5) and multiplying through by p_i^T to give:

$$\Xi p_i^{\mathrm{T}} = t_1 p_1 p_i^{\mathrm{T}} + t_2 p_2 p_i^{\mathrm{T}} + \ldots + t_N p_N p_i^{\mathrm{T}} + E_0 p_i^{\mathrm{T}} \qquad (6)$$

given that $t_i p_i p_j^{\mathrm{T}} = 0 \quad (\forall i : i \neq j)$, and $E_0 p_i^{\mathrm{T}} = 0$, equation (6) reduces to:

$$\Xi p_i^{\mathrm{T}} = t_1 p_1 p_i^{\mathrm{T}} \qquad (7)$$

$$\therefore t_1 = \frac{\Xi p_i^{\mathrm{T}}}{p_1 p_i^{\mathrm{T}}} \qquad (8)$$

$t_2 \ldots t_n$ can be similarly computed.

Selection of principal components from equation (6), by taking the first m terms reduces the data set to that, for example, of equation (9). The value of m will determine the accuracy of the simplification. If m=2, then:

$$\Xi = t_1 p_1 + t_2 p_2 + E_1 \qquad (9)$$

where E_1 is an error signal and all other terms take values as hereinebefore described. The dimension of E_1 will be equal to the dimension of the input data set Ξ.

In accordance with the flow diagram of figure 3, the error signal E is fed as input to the non-linear principal component analysis module 15. The purpose of the non-linear principal component analysis module 15 is to convert the error signal into a plurality of non-linear principal components, thereby reducing the dimension of the error signal E_1. This is achieved by considering an equation of the form:

$$E_1 = f(t_3, t_4) \qquad (10)$$

where:

f is a non-linear function whose form may be determined using a neural network; and

t_3 and t_4 are non-linear principal components whose form is determined using a neural network.

Determination of f, t_3, t_4 will now be described. Given $t_1 p_1$ and $t_2 p_2$ from equation (9), a matrix X may be constructed such that:

$$X = \Xi - t_1 p_1 - t_2 p_2 \tag{11}$$

This matrix X is equivalent to the error signal E and is input to the non-linear principal component analysis module 15. It can be seen from figure 3 that the non-linear module 15 uses two neural networks. The structure of these two ANN's is illustrated in figure 4.

Referring to figure 4, a first ANN comprises three layers 17, 18, 19 and a second ANN comprises a layer 19 in common with the first ANN together with two further layers 20, 21.

The first ANN is considered to correspond to the function f^{-1}, while the second is considered to correspond to the function f, where f^{-1} is the inverse of f. The output of the first ANN f^{-1} is the input to the second ANN f and this output comprises t_3, t_4. The relationship is as represented in equation (10).

Since f^{-1} is the inverse of f, the input X and output $\hat{X}$ should be equal. Given this information, and a known form of ANN, the weights of each neuron of the ANN may be determined so as to calculate f, and t_3, t_4 with sufficient accuracy by means of a training algorithm such as that referred to above.

The first and second ANNs each comprise one or more layers of neurons. When creating suitable ANNs, it is possible to initially use a small number of neurons, and to increase this number in both the first and second ANNs concurrently. After each

neuron addition, the output of the second ANN is compared to the input to the first ANN and the increase in neuron number is continued until the comparison shows that a difference between values being compared is below a predetermined threshold. Each of the first and second ANNs has an equal number of neurons. It has been found that creating ANN's where the central (or hidden) layers 18, 20 of each ANN have twice as many neurons as the input layer 14 of the first neural network gives satisfactory results.

The network of figure 4 is a well known ANN structure generally referred to as a feed-forward back propagation network. The central layer 19 of the structure, determines the principal components and thus the number of neurons in this layer determines the efficiency of the data reduction provided.

Configuring the ANN such that the outputs and inputs $X, \hat{X}$ are substantially equal is carried out using an auto correlation technique. The input data X is a matrix having equal dimension to the input data applied to the linear principal component alanysis module 14 and is made up of a plurality of row vectors, each row vector containing temporally spaced values for a particular measurement. Each row represents a different measurement such that the matrix has as many rows as there are measurements.

Auto correlation is performed in a conventional way, such that differences between the matrix $\hat{X}$ and the matrix X are compared. The result of the auto correlation test is presented in figure 5. The *x-axis* represents shift, and the *y-axis* represents the quality of correlation at that shift. Therefore, figure 5 shows a good correlation at shift zero, as would be expected. A large central peak (height A), which is reasonably narrow (width B), and small peaks of maximum height C show that the differences between $\hat{X}$ and X are small and are mainly made up of independent noise (also known as white noise), that can not be modelled. However, the illustration of figure 5 shows that there is little white noise, denoted by smaller peaks (i.e. correlation at non-zero shift values), thereby further indicating that a good correlation has been obtained.

The auto correlation process can be conveniently carried out using MATLAB® (distributed by The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098 United States of America), using an "xcorr" function supplied therein. The correlation data shown in figure 5 is generated by MATLAB, and only data in a specified auto correlation area is used in the correlation calculations plotted in figure 5. This area is defined as that between the second and third maxima in the data set.

The effectiveness of the correlation illustrated in figure 5 is mathematically represented by a factor F given in equation (12):

$$F = \frac{Width * SecondMaxPeak}{MaxPeak} \qquad (12)$$

where,

F is a Factor of the auto correlation in the specified area;

Width is a distance (along the x axis)between the two lowest values within the illustrated area;

SecondMaxPeak is the second highest peak within the area; and

MaxPeak is the height of the maximum peak (A in figure 5)

The width of the specified area auto correlation is the distance between the two smallest values within the plotted area. Width can therefore by computed according to equation (13)

$$Width = |POS_1 - POS_2| \qquad (13)$$

where:

POS1 is the position (on the *x-axis* of figure 4) of the minimum value; and

POS2 is the position (on the *x-axis* of figure 4)of the second lowest value.

The highest peak is calculated by subtracting the maximum and minimum values plotted in the specified area.

$$MaxPeak = \max(VAL) - \min(VAL) \qquad (14)$$

where,

VAL is a vector of all values plotted in figure 5;

max() is a predefined function taking a vector, and returning the maximum value contained within that vector; and

min() is a predefined function taking a vector, and returning the minimumvalue contained within that vector; and

$$SecondMaxPeak = VAL(1,2) - \min(VAL) \qquad (15)$$

where,

$VAL(1,2)$ is the second highest value of vector VAL; and

min is a predefined function as defined above.

The factor F calculated using equation (12) is a measure of the accuracy of the correlation. In a preferred embodiment of the present invention, all values input to the module 15 (that is input to layer 17 of figure 4) are normalised (that is take values between -1 and $+1$). In such a circumstance it has been found that training the neural network until F has a value less than 1 provides good results. Further training is likely to improve the accuracy of results obtained, but it is likely that such training will result in relatively modest improvements.

Having performed both the linear and non-linear principal component analysis techniques described above, the linear and non-linear principal components t_1, t_2, t_3, and t_4, can be used to represent the input dataset. Thus, the dimension of the input data has been reduced considerably.

Referring back to figure 3, it can be seen that the non-linear principal components t_3,

and t_4 are taken from the data output from the first ANN f^1. These, together with the linear principal components t_1, and t_2, are used as input to the model 16. The model 16 is recursive, taking previous output values as input as illustrated in figure 3. This allows the model to be dynamic in nature, as is explained below.

The model 16 is represented by at least one further ANN having a form as set out in equation (1). At least one output relating to the input values t_1, t_2, t_3, and t_4 is predicted by the model 16. In use, measurements will be taken and used to generate input into the model to predict an output. This output should match at least one predetermined property of the manufactured paper, for example paper strength.

This problem is complicated because generally the model 16 will incorporate a number of ANNs will be operating concurrently, each having the ability to optimise one criteria of the paper machine output. These ANNs must all be optimised concurrently, so as to obtain input parameters to create output which accurately reflects all desired properties. Additionally, it is desirable to optimise chemical input quantities so as to control costs. A method of performing this optimisation will now be described in terms of a system where output strength and chemical retention are both to be optimised.

Suitable equations for each of these characteristics are:

$$S_k = \breve{f}(S_{k-1}, S_{k-2}, \ldots, S_{k-l}, R_{k-1}, R_{k-2}, \ldots, R_{k-l}, t_{1(k-1)}, t_{2,(k-1)}, t_{3(k-1)}, t_{4(k-1)}, U_{(k-d)}) \qquad (16)$$

$$R_k = \breve{g}(S_{k-1}, S_{k-2}, \ldots, S_{k-l}, R_{k-1}, R_{k-2}, \ldots, R_{k-l}, t_{1(k-1)}, t_{2,(k-1)}, t_{3(k-1)}, t_{4(k-1)}, U_{(k-d)}) \qquad (17)$$

where

S represents the strength of output paper;

R represents retention of raw materials in the output paper;

g and f are suitably trained ANNs;

k is the current sample;

k-1 is the previous sample;

$t_1,\ldots,t_4$ are as hereinbefore defined;

$U_{(k-d)}$ is a variable related to chemical costs;

d is time delay; and

l is the time period over which previous measurements are to be taken into account.

Equations (16) and (17) provide a dynamic model by linking values at a time k to values at a time k-1. The presence of R and S on the right hand side of the equation provides a feedback loop, such that previous output values affect the current output value.

It is desired to use the system of the present invention so as to predict output properties using currently known values. The equations may be modified to do this such that:

$$S_{k+d} = \breve{f}(S_k, S_{k-1}, S_{k-2},\ldots, S_{k-(l-1)}, R_k, R_{k-1}, R_{k-2},\ldots, R_{k-(l-1)}, t_{1(k)}, t_{2,(k)}, t_{3(k)}, t_{4(k)}, U_{(k)}) \qquad (18)$$

$$R_{k+d} = \breve{g}(S_k, S_{k-1}, S_{k-2},\ldots, S_{k-(l-1)}, R_k, R_{k-1}, R_{k-2},\ldots, R_{k-(l-1)}, t_{1(k)}, t_{2,(k)}, t_{3(k)}, t_{4(k)}, U_{(k)}) \qquad (19)$$

In order to ensure that equations (18) and (19) accurately reflect measured properties of the manufactured paper, modelled and measured values should be compared. This can conveniently be done using an auto correlation technique similar to that described above. If the measured and modelled values are not sufficiently close, amendments to the model are necessary.

In a paper manufacturing plant, a number of dynamic models having the form of equations (18) and (19) are created. It is necessary to combine the models of equations (18) and (19) so as to obtain a performance function made up of the two models. A number of suitable equations have been developed, and are detailed in equations (20) to (22) below.

$$J_1 = \frac{a}{S_{k+d}} + \frac{b}{R_{k+d}} + cU_k^{\ 2} \tag{20}$$

$$J_2 = a(S_{k+d} - T)^2 + \frac{b}{R_{k+d}} + cU_k^{\ 2} \tag{21}$$

$$J_3 = a(S_{k+d} - T)^2 + b(R_{k+d} - T_r)^2 + cU_k^{\ 2} \tag{22}$$

where:

J_1, J_2, J_3 are variables representing performance;

a, b and c are constants; and

T and T_r are target values.

Referring to equation (20) it can be seen that minimising J_1 will result in maximising both R_{k+d} and S_{k+d}, while minimising U. As both Strength *(S)* and Retention *(R)* should be maximised, and Chemical Inputs *(U)* should be minimised, it can be seen that J_1 is effective as a performance function.

Equation (21) is such that its first term $(S_{k+d} - T)^2$ will be a minimum when $S_{k+d} = T$. Thus, minimising J_2 will result in maximising of Retention *(R)*, minimising of chemical inputs, and focusing Strength *(S)* on its target value.

Equation (22) provides a performance function which can be minimised to allow both strength and retention to be focussed upon predetermined targets.

Referring back to equations (18) and (19), it will be seen that each of $S_k, R_k, t_{1(k)}, t_{2,(k)}, t_{3(k)}, t_{4(k)}$ are known; $t_{1(k)}, t_{2,(k)}, t_{3(k)}, t_{4(k)}$ having been obtained as output from the data reduction method and S_k, R_k, being obtainable through measurement. The value of U_k is also known.

The only directly variable parameter is U_k. An initial value is obtained by carrying out data analysis based upon past performance. If this does not yield desired values for J_i in any of equations (20), (21) or (22), U_k is varied, until it reaches an optimum value, which may be determined by:

$$U_{k+d} = U_k - A\frac{\delta J}{\delta U} \tag{23}$$

where A is a constant.

It will be appreciated that similar principles may be applied to the creation and optimisation of a performance function made up of more than two models, including models representing qualities such as breaking frequency, effluent quality, and energy consumption. Equation (24) provides an example of such as model:

$$J_4 = a(S_{k+d} - T)^2 + b(R_{k+d} - T_r)^2 + Bf + Eq + Ec + cU_k^{\ 2} \tag{24}$$

where all terms are as hereinbefore defined and:

J_4 is a performance function;

Bf is a term involving a model representing breaking frequency;

Eq is a term involving a model representing Effluent Quality; and

Ec is a term involving a model representing Energy consumption.

CLAIMS

1. A method for converting a quantity of input data to output data of lesser dimension, comprising:

applying a linear principal component analysis technique to the input data to generate a plurality of linear principal components and an error signal;

inputting said error signal to a first neural network which outputs at least one variable;

inputting the said at least one variable to a second neural network; and

configuring the first and second neural networks such that the output of the second neural network is substantially equal to the input to the first neural network;

the output data being represented by the said plurality of linear principal components and the said at least one variable.

2. A method according to claim 1, wherein the first and second neural networks are trained until the difference between the output of the second neural network and the input to the first neural network is less than a predetermined threshold.

3. A method according to claim 1 or 2, wherein neurons are added to the first and second neural networks until the difference between the output of the second neural network and the input to the first neural network is less than a predetermined threshold.

4. A method according to any preceding claim, wherein a difference between the input to the first neural network and the output of the second neural network is assessed using an auto correlation technique.

5. A method according to any preceding claim, wherein the first and second neural networks each comprise an input layer having connections with a hidden layer which in turn has connections with an output layer, the output layer of the first neural network forming the input layer of the second neural network.

6. A method according to claim 5, wherein the hidden layer of each neural network has more neurons than the input layer of the first neural network.

7. A method according to claim 6, wherein the hidden layer of each neural network has twice as many neurons as the input layer of the first neural network.

8. A method according to claim 5, 6 or 7, wherein the input layer of the first neural network has an equal number of neurons to the output layer of the second neural network.

9. A method of dynamically modelling a paper manufacturing plant comprising derivation of a function which takes as input a plurality of parameters of said paper manufacturing plant which have been reduced in dimension and outputs a value indicative of a quality of paper output from the paper manufacturing plant.

10. A method according to claim 9, wherein the plurality of parameters comprise a plurality of temporally spaced values for at least one parameter of the paper manufacturing plant.

11. A method according to claim 9 or 10, wherein the function also takes as input values previously output from the function.

12. A method according to claim 9, 10 or 11, wherein the function is provided by an artificial neural network.

13. A method according to claim 9, 10, 11 or 12, wherein at least some of the plurality of parameters are outputs from a method according to any one of claims 1 to 8.

14. A method according to any one of claims 9 to 13, wherein the model has a general form:

$$R_{k+d} = \breve{g}(S_k, S_{k-1}, S_{k-2}, \ldots, S_{k-(l-1)}, R_k, R_{k-1}, R_{k-2}, \ldots, R_{k-(l-1)}, t_{1(k)}, t_{2,(k)}, t_{3(k)}, t_{4(k)}, U_{(k)})$$

where: R represents a quality of the output paper;

k is a current sample;

k-1 is a previous sample;

$t_1, \ldots t_4$ are parameters of the paper manufacturing plant

$U_{(k-d)}$ is a variable related to chemical costs;

d is time delay; and

l is the time period over which previous measurements are to be taken into account.

15. A method wherein the outputs of a plurality of models generated using the method of any one of claims 9 to 14 are combined to generate a performance function.

16. A method according to claim 15, wherein optimisation of the performance function seeks to maximise at least one of the outputs of the plurality of models.

17. A method according to claim 15, wherein optimisation of the performance function seeks to focus at least one of the outputs of the plurality of models upon a predetermined target.

18. A method according to claim 15 wherein the performance function has the form:

$$J = \frac{a}{A} + \frac{b}{B} + C$$

where: J is an output from the function;

A and B are outputs from models generated using a method according to any one of claims 1 to 8;

C is a variable representing all adjustable inputs that are to be minimised; and

a and b are constants.

19. A method according to claim 15, wherein the performance function has the form:

$$J = a(A-T)^2 + \frac{b}{B} + C$$

where: J is an output from the function;

A and B are outputs from models generated using a method according to any one of claims 1 to 8;

C is a variable representing all adjustable inputs that are to be minimised;

T is a target upon which A is to be focussed; and

a and b are constants.

20. A method according to claim 15, wherein the performance function has the form:

$$J = a(A-T)^2 + b(B-T_2)^2 + C$$

where: J is an output from the function;

A and B are outputs from models generated using a method according to any one of claims 1 to 10;

T and T_2 are target values for A and B respectively; and

a and b are constants.

21. A method of optimising operation of a paper manufacturing plant, comprising maximising a performance function generated using a method according to any one of claims 15 to 20.

22. A computer program for carrying out the method of any preceding claim.

23. A data carrier medium carrying computer program code means to cause a computer to execute procedure in accordance with the method of any one of claims 1 to 21.

24. An apparatus for carrying out a method according to any one of claims 1 to 21.

25. A method substantially as hereinbefore described with reference to figures 3 to 5 of the accompanying drawings.

26. An apparatus for carrying out a method substantially as hereinbefore described with reference to figures 3 to 5 of the accompanying drawings.

27. A computer program for carrying out a method substantially as hereinbefore described, with reference to figures 3 to 5 of the accompanying drawings.

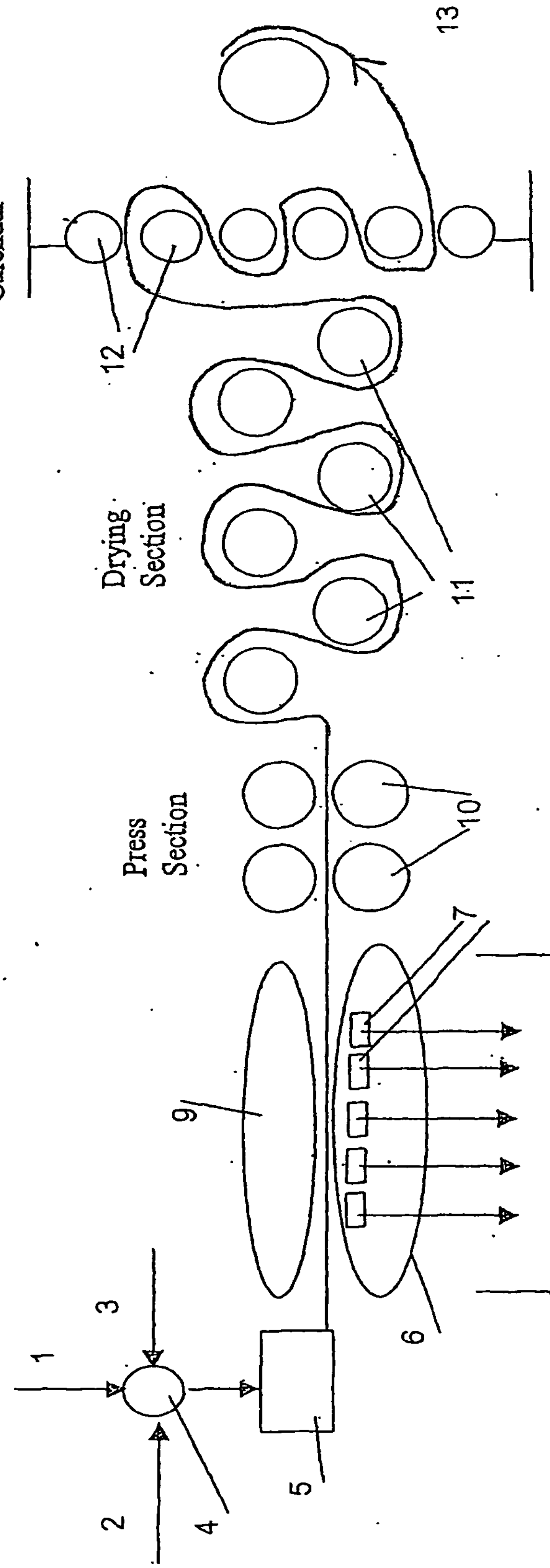

Calendar
Drying Section
Press Section
1
2
3
4
5
6
7
8
9
10
10
11
12
13

FIG.1

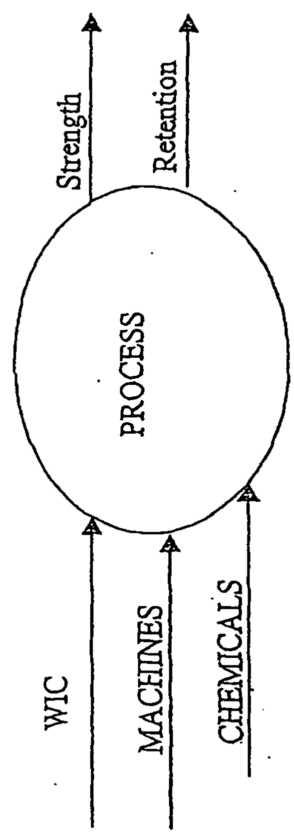

Strength
Retention
PROCESS
WIC
MACHINES
CHEMICALS

FIG.2

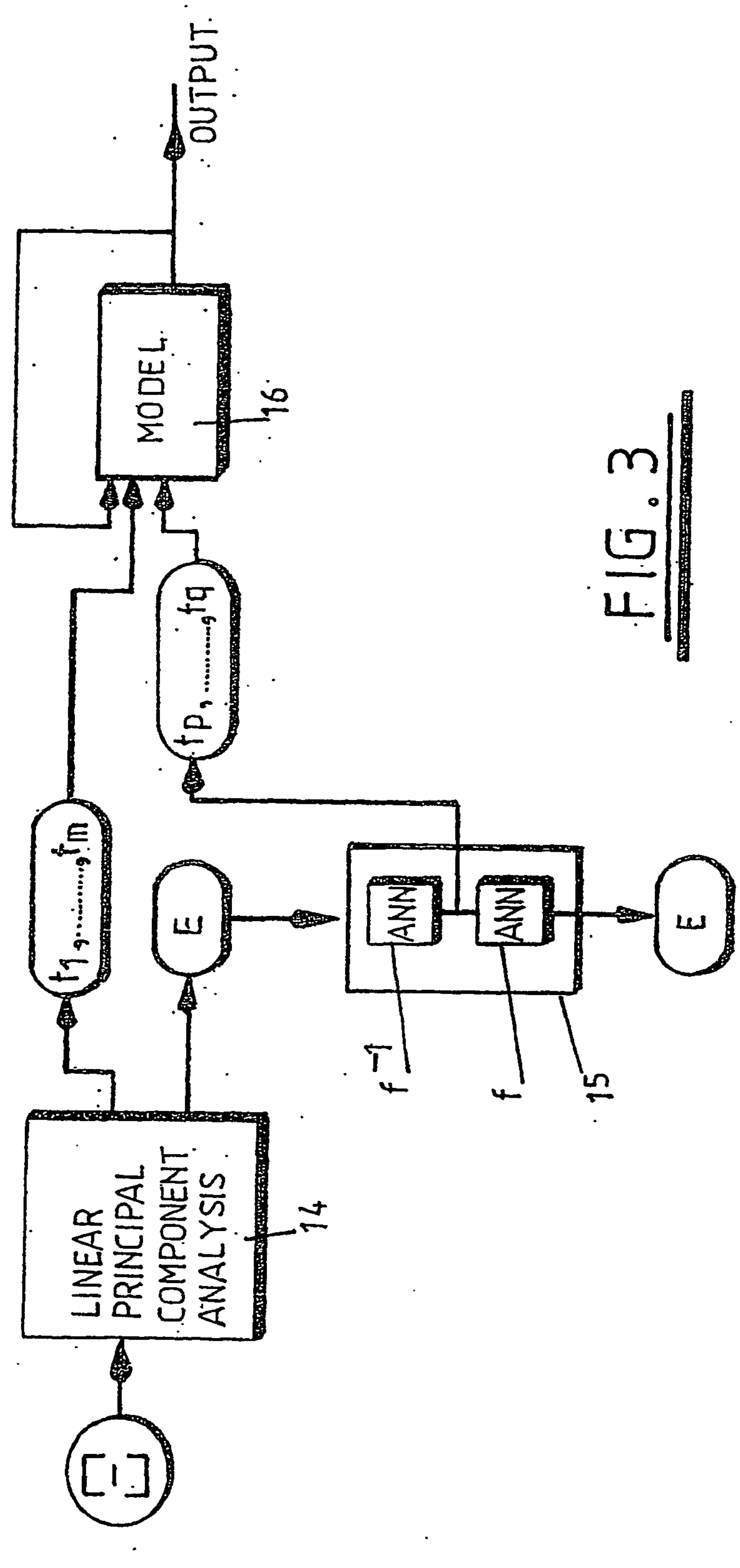

OUTPUT
MODEL
16
LINEAR PRINCIPAL COMPONENT ANALYSIS
14
t1,.........,tm
tp,.........,tq
E
ANN
ANN
E
f⁻¹
f
15

FIG. 3

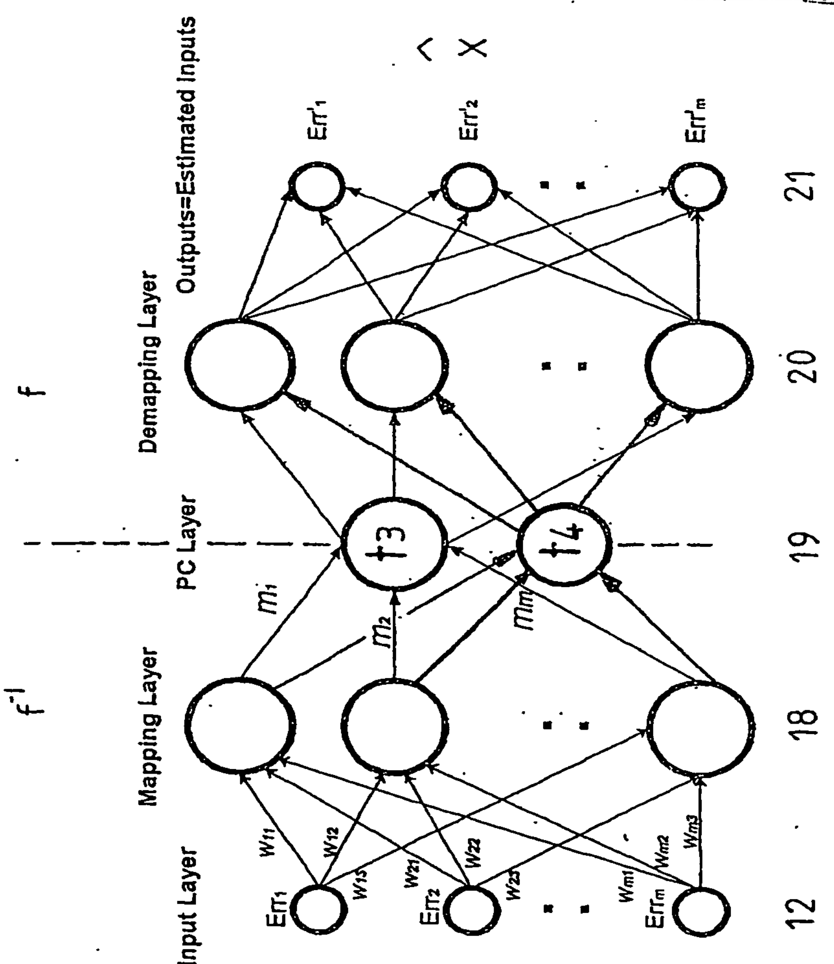

Input Layer
Mapping Layer
PC Layer
Demapping Layer
Outputs=Estimated Inputs
f
f
Err1
Err2
Errm
W11
W12
W13
W21
W22
W23
Wm1
Wm2
Wm3
m1
m2
mm
t3
t4
Err'1
Err'2
Err'm
X
12
18
19
20
21

FIG.4

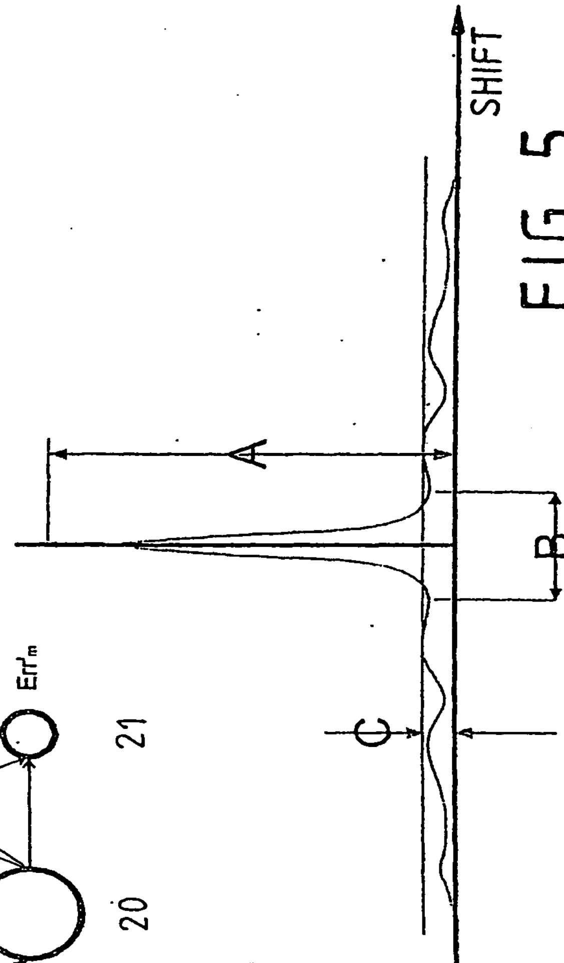

A
B
C
SHIFT

FIG.5

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 03/00943

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 D21G9/00 G05B13/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 D21G G05B

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	US 6 405 140 B1 (BONISSONE PIERO PATRONE ET AL) 11 June 2002 (2002-06-11) column 1, line 15 -column 2, line 2 column 3, line 45 -column 4, line 55 column 5, line 10 - line 48 column 13, line 38 - line 67 abstract; figures 2-4,6,12,18 ---	9-24
Y	US 6 212 509 B1 (MENG ZHUO ET AL) 3 April 2001 (2001-04-03) column 2, line 11 - line 64 column 3, line 36 - line 45 column 3, line 56 -column 4, line 4 column 4, line 15 - line 20 column 4, line 63 -column 5, line 16 column 5, line 29 - line 56 abstract; figures 1A,1B,14 --- -/--	9-24

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 June 2003	24 07 2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	ISMAR HADZIEFENDIC/JA

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 03/00943

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2001/028459 A1 (HARTENSTEIN HERMANN ET AL) 11 October 2001 (2001-10-11) paragraph [0001] - paragraph [0003] paragraph [0013] - paragraph [0017] paragraph [0022] - paragraph [0023] paragraph [0028] paragraph [0042] - paragraph [0043] paragraph [0045] - paragraph [0047]; figure 2 ---	9-24
A	US 6 187 145 B1 (ZEINER GERHARD ET AL) 13 February 2001 (2001-02-13) column 1, line 1 -column 2, line 62 column 4, line 35 - line 53 abstract; claims 1,10; figures 1-7 ---	9-24
A	NORIEGA J R ET AL: "A direct adaptive neural-network control for unknown nonlinear systems and its application" IEEE TRANSACTIONS ON NEURAL NETWORKS, JAN. 1998, IEEE, USA, vol. 9, no. 1, pages 27-34, XP002244890 ISSN: 1045-9227 abstract; figures 1-3,9 ---	9-24
A	US 5 805 453 A (SASAKI TAKASHI) 8 September 1998 (1998-09-08) column 1, line 12 - line 62 abstract; figures 1-3 ---	9-24
A	US 5 841 671 A (FURUMOTO HERBERT) 24 November 1998 (1998-11-24) page 1, line 1 -page 5, column 7 abstract; claims 1-4; figure 1 -----	9-24

INTERNATIONAL SEARCH REPORT

International application No.
PCT/GB 03/00943

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 1-8
because they relate to subject matter not required to be searched by this Authority, namely:

The subject matter according to claims 1-8 describes a mathematical calculation method using only generally known technical tools (see Article 39 of the PCT).

2. [X] Claims Nos.: 25-27
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/GB 03/00943

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 25-27

The subject matter defined by claims 25-27, is only described by reference to drawings which according to Rule 6.2(a) is to be avoided if not absolutely necessary.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

ation on patent family members

International Application No
PCT/GB 03/00943

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6405140 B1	11-06-2002	US 2002038197 A1	28-03-2002
US 6212509 B1	03-04-2001	US 6134537 A	17-10-2000
		US 5734796 A	31-03-1998
		US 2001032198 A1	18-10-2001
		AU 737276 B2	16-08-2001
		AU 1912699 A	05-07-1999
		BR 9813585 A	26-12-2001
		CA 2312902 A1	24-06-1999
		CN 1282435 T	31-01-2001
		EP 1038261 A1	27-09-2000
		JP 2002509303 T	26-03-2002
		WO 9931624 A1	24-06-1999
US 2001028459 A1	11-10-2001	DE 19850825 A1	18-05-2000
		AT 235603 T	15-04-2003
		CA 2350559 A1	11-05-2000
		WO 0026610 A2	11-05-2000
		DE 59904779 D1	30-04-2003
		EP 1127189 A2	29-08-2001
US 6187145 B1	13-02-2001	DE 19510009 A1	26-09-1996
		AT 180848 T	15-06-1999
		BR 9607838 A	16-06-1998
		CA 2215944 A1	26-09-1996
		WO 9629468 A1	26-09-1996
		DE 59602093 D1	08-07-1999
		EP 0815320 A1	07-01-1998
		NO 974357 A	21-11-1997
		ZA 9602316 A	27-09-1996
US 5805453 A	08-09-1998	JP 10025688 A	27-01-1998
US 5841671 A	24-11-1998	AT 150114 T	15-03-1997
		BR 9407546 A	31-12-1996
		CN 1131445 A ,B	18-09-1996
		DE 59402103 D1	17-04-1997
		WO 9508019 A1	23-03-1995
		EP 0719358 A1	03-07-1996
		ES 2098980 T3	01-05-1997
		FI 961235 A	15-03-1996
		NO 961066 A	18-03-1996